**7PAM2000**

**APPLIED DATA SCIENCE 1**

ASSIGNMENT 1: VISUALISATION

BY

FULLNAME: KAMALI BAKTHAVATCHALAM

STUDENT ID: 22081557

MAIL ID: [kb23aat@herts.ac.uk](mailto:kb23aat@herts.ac.uk)

DATA SOURSE: https://www.kaggle.com/code/sagar2439/walmart-store-sales-prediction-a-simple-approach/input

INTRODUCTION:

In this assignment we are getting to know about the line plot, graph plot, pie chart and bar chart from the data on walmart store sales prediction with the functions and data source and using matplot , pandas .

1.Produce a line plot showing multiple lines with proper labels and legend. Describe what conclusions you can draw from this plot.

VISUALISATION1: FUELPRICE VS UNEMPLOYMENT (LINE PLOT)

A graph showing a graph of a graph

Description automatically generated with medium confidence

EXPLANATION:

• The multi-line chart denotes the relationship between Stores and the factors such as Unemployment and Fuel Price that have an impact on the customer flow. The orange line denotes the unemployment and blue line denotes fuel prices. The unemployment has more variations with respect to store compared to fuel prices.

•The chart denotes the comparison of stores with various factors.

•The factors taken into consideration are Fuel price and Unemployment.

2.Produce graphs using two other visualisation methods. Explain why you picked this type of graph and describe what conclusions you can draw. Marks will be awarded for plausible choices of graphs. The choice does not need to be the perfect one, but at least one type will not a good choice if you visualise exactly the same data three times. Note, that we consider it not exactly the same, if you do a line plot of a time series and then use a pie chart to visualise relative sizes for a selected time.

VISUALISATION 2: HOLIDAY & NON-HOLIDAY COUNT ( PIE CHART)

A blue circle with orange triangle and text

Description automatically generated

EXPLANATION:

• The pie chart represents the customer flow in Holidays and Non- Holidays. It is created pyplot function from matplotlib. The colour representation is default as per the visual. The labels given are holiday and non-holiday respectively for 0’s and 1’s. The chart shows that the customer flow is comparatively more during the Holidays than that of the non-holidays.

• The pie chart denotes the count of purchase on holidays and non-holidays

•The purchase on holidays is more than non-holidays.

VISUALISATION 3: AVERAGE SALE ON NON -HOLIDAY AND SPECIAL HOLIDAY (BAR CHART)

A graph of blue squares

Description automatically generated

EXPLANATION:

• The bar chart denotes the average sales that has taken place in both non-holiday week and holiday week. The average sales are calculated by considering the weekly sales based on the holiday flag mentioned in the data. The chart denotes that special holiday week has more sales compared to non-holiday week.

•The chart denotes the comparison of average sales of holiday week and non-holiday week.

•It is observed that the average sales of holiday week is more compared to non-holiday week.

**ANOTHER EXAMPLE FOR FUEL PRICE VS UNEMPLOYMENT.**

VISUALISATION 4: FUEL PRICE VS UNEMPLOYMENT (LINE CHART)

A graph of a line

Description automatically generated with medium confidence

EXPLANATION:

• The multi-line chart denotes linear relationships between the holiday flag and factors such as fuel price and unemployment. The blue line denotes Fuel Price and orange line denotes the unemployment rate, which shows a peak in both the factors, when there is a holiday and decline. Whereas,

•The orange colour line denotes the variation of Unemployment with respect to Holiday Flag

•The blue colour line denotes the variation of Unemployment with respect to Holiday Flag.

REFERENCE:

MATPLOT: I learned matplot from my lecture class

PANDAS: I learned pandas by lecture class

DATA SOURSE: https://www.kaggle.com/code/sagar2439/walmart-store-sales-prediction-a-simple-approach/input